CLAIMS

That which is claimed:

1.   A method, comprising:

defining target rules for detection of target hits in an article, including defining

a target article region;

defining extraction rules based on the target rules for the extraction of extracts

from the article, including an extraction article region;

applying target rules to each target article region of the article to determine

target hits; and

applying extraction rules to detect at least one extract from the article based on

the determined target hit.

2.   The method of claim 1, wherein a plurality of target hits are detected and a

plurality of extracts are extracted.

3.   The method of claim 2, further comprising sorting the extracts based on the

extraction rules.

4.   The method of claim 1, further comprising extracting at least one extract from

the article based on the determined target hit.

5.   The method of claim 1, wherein the target rules further comprise a target definition and a target score formula.

6.   The method of claim 1, wherein applying the target rules comprises using the target score formula to detect target hits.

7.   The method of claim 5, wherein the target definition comprises a concept set, a gist or both.

8.   The method of claim 7, wherein the concept set is a list of concepts.

9.   The method of claim 8, wherein concepts in the list of concepts are produced by set operations on multiple lists of concepts.

10.   The method of claim 7, wherein a gist comprises weighted concepts.

11.   The method of claim 7, wherein the gist is user defined or is a calculated gist of the article.

12.   The method of claim 1, wherein the target article region is a article, a sentence or a term.

13.   The method of claim 1, wherein the extraction article region is a article, a sentence or a term.

14.   The method of claim 1, wherein the article is preprocessed to determine concepts contained in the article and a gist for the article.

15.   A computer-readable medium containing program code, comprising:

program code for defining target rules for detection of target hits in an article, including defining a target article region;

program code for defining extraction rules based on the target rules for the extraction of extracts from the article, including an extraction article region;

program code for applying target rules to each target article region of the article to determine target hits; and

program code for applying extraction rules to detect at least one extract from the article based on the determined target hit.

16.   The computer-readable medium of claim 15, wherein a plurality of target hits are detected and a plurality of extracts are extracted.

17.   The computer-readable medium of claim 16, further comprising program code for sorting the extracts based on the extraction rules.

18.   The computer-readable medium of claim 15, further comprising program code for extracting at least one extract from the article based on the determined target hit.

19.   The computer-readable medium of claim 15, wherein the target rules further comprise a target definition and a target score formula.

20.   The computer-readable medium of claim 15, wherein applying the target rules comprises using the target score formula to detect target hits.

21.   The computer-readable medium of claim 19, wherein the target definition comprises a concept set, a gist or both.

22.   The computer-readable medium of claim 21, wherein the concept set is a list of concepts.

23.   The computer-readable medium of claim 22, wherein concepts in the list of concepts are produced by set operations on multiple lists of concepts.

24.   The computer-readable medium of claim 21, wherein a gist comprises weighted concepts.

25

25.   The computer-readable medium of claim 21, wherein the gist is user defined or is a calculated gist of the article.

26.   The computer-readable medium of claim 15, wherein the target article region is a article, a sentence or a term.

27.   The computer-readable medium of claim 15, wherein the extraction article region is a article, a sentence or a term.

28.   The computer-readable medium of claim 15, wherein the article is preprocessed to determine concepts contained in the article and a gist for the article.